Geometrical-Based Lip-Reading Using Template Probabilistic Multi-Dimension Dynamic Time Warping

M.Z. Ibrahim[a] , D.J. Mulvaney[b]

[a] Faculty of Electrical and Electronics Engineering, University Malaysia Pahang, 26300 Pahang, Malaysia

[b] School of Electronic, Electrical and Systems Engineering, Loughborough University, LE11 3TU, United Kingdom

ABSTRACT

By identifying lip movements and characterizing their associations with speech sounds, the performance of speech recognition systems can be improved, particularly when operating in noisy environments. In this paper, we present a geometrical-based automatic lip reading system that extracts the lip region from images using conventional techniques, but the contour itself is extracted using a novel application of a combination of border following and convex hull approaches. Classification is carried out using an enhanced dynamic time warping technique that has the ability to operate in multiple dimensions and a template probability technique that is able to compensate for differences in the way words are uttered in the training set. The performance of the new system has been assessed in recognition of the English digits 0 to 9 as available in the CUAVE database. The experimental results obtained from the new approach compared favorably with those of existing lip reading approaches, achieving a word recognition accuracy of up to 71% with the visual information being obtained from estimates of lip height, width and their ratio.

KEYWORDS: Lip reading; Lip geometry; Mouth detection; Skin segmentation; Convex hull; Multi dimension dynamic time warping; Template probabilistic; OpenCV

DOI: 10.1016/j.jvcir.2015.04.013